**1007 Project Proposal: NYC Motor Vehicle Collisions Visualization**

Name / NetID:

Li Lin Qin llq205

Xue Yang xy990

Yidi Zhang yz3464

Being struck by a vehicle is the second-leading cause of injury-related death for people. On average, vehicles seriously injure or kill a New Yorker every two hours. For the term project, we will present the data visualization of vehicle collisions in New York City, using the dataset provided by the Police Department. Our goal is to promote public awareness, and hopefully, encourage people’s contribution to street safety.

Here is a link for our dataset: <https://data.cityofnewyork.us/Public-Safety/NYPD-Motor-Vehicle-Collisions/h9gi-nx95>

There are about total of 915044 observations, dating from 07/01/2012 till 10/2016.

Variables provided include: Date, Time, Borough, Zip Code, Location, Number of Persons Injured / Killed, Vehicle Types.

1. In general, there will be a Geomap of all vehicle collisions in the past few years at NYC
2. Along with the geomap, a pie chart of motor collisions according to different boroughs will be presented.
3. There will be a filter area that includes zip code, borough, date(year, month, day), vehicle type, in which users can input any appropriate values, then there will be some figure outputs according to their inputs (specify below).

|  |  |
| --- | --- |
| **User Input** | **Presented Output** |
| Zip code (NYC zip code) | Corresponding collisions in the area light up in the map; with details pop-up when moving over |
| Borough (5 boroughs in NYC) | Geomap of collision in this borough; color saturation indicating different injury rates |
| Year (From 2012 to 2016) | Matplotlib pop-up with collisions over the year |
| Month | Matplotlib pop-up with collisions over the month |
| Date | Matplotlib pop-up with collisions over the day |
| Vehicle Type | Geomap of this kind of vehicle collision, with histogram comparing with other vehicle types |

How our project will utilize:

Pandas + NumPy: load data, clean data, preliminary analysis of data

Matplotlib: displaying results using different graphs, including pie chart, histogram, line plots...